MODIS Aerosol Optical Depth Bias Adjustment Using Machine Learning Algorithms

AGU 2011 – Poster A53C-0371

Arif Albayrak1, Jennifer Wei2, Maksym Petrenko3, David Lary4, and Gregory Leptoukh5

1,2 - NASA GES DISC, ADNET , 3 – NASA, UMCP, 4 - University of Texas at Dallas, 5 – NASA GES DISC

Goal and Application

Goal : The goal of this work is to adjust the biases and the systematic errors in the MODIS (both Terra and Aqua) aerosol product, using Machine Learning algorithms (i.e., Neural Network).

Application : The results of bias adjustment for MODIS Terra and Aqua are incorporated into the AeroStat Giovanni as part of the NASA ACCESS funded AeroStat project (see Poster IN51C-1604).

Background

Over the past decade, global aerosol observations have been conducted by space-borne sensors, airborne instruments, and ground-base network measurements. Unfortunately, quite often we encounter the differences of aerosol measurements by different well-calibrated instruments, even with a careful collocation in time and space. The differences might be rather substantial, and need to be better understood and accounted for when merging data from many sensors. The possible causes for these differences come from instrumental bias, different satellite viewing geometries, calibration issues, dynamically changing atmospheric and the surface conditions, and other "regressors", resulting in random and systematic errors in the final aerosol products.

Methodology

Neural Networks: Neural Networks (NN) to remove biases and the systematic errors from MODIS (both Terra and Aqua) aerosol product. The Aerosol Robotic Network of sun-photometers (AERONET) are used as a baseline for evaluating the MODIS aerosol products. The overall NN algorithm is described as following:

- Network Architecture: Feed-Forward (no cycles)
- Optimization Algorithm: Back Propagation
- Hidden Layers: One
- Nodes: 40
- Number of Regressors
  - Dark Target Land: 15
  - Dark Target Ocean: 14

NOTE: While it may be controversial to have "optimal sets of regressors" in our system when comparing global aerosol products, there is a parallel study on this topic by one of our co-authors, David Lary.

Code: Python and ffnet module by Merek Wojciechowski (Fortran + f2py)

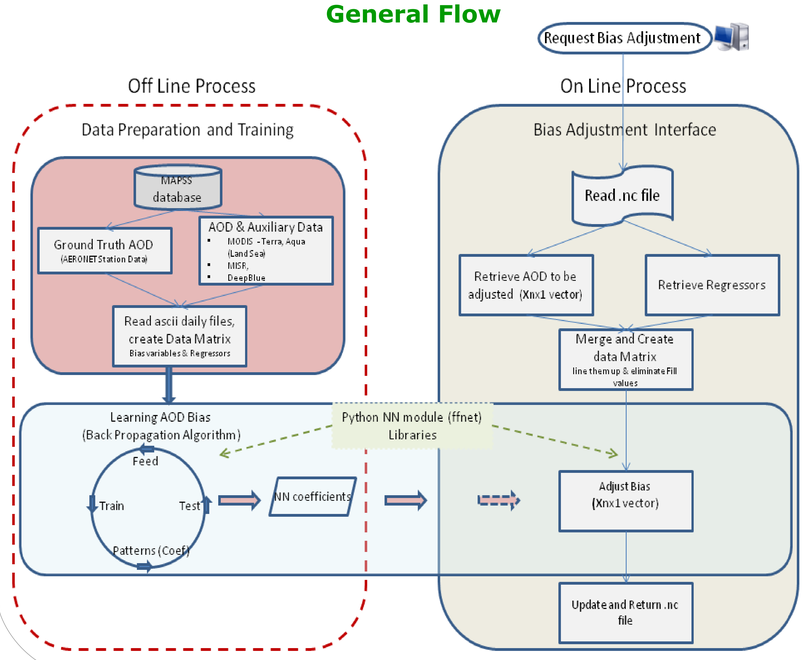
Statistical Methods and Visualization Tools:
Because of high numbers of data with different regimes, scatter plot with 1:1 line is not enough to draw conclusions. Results are further supported with

- Standard and Non-standard tests for analysis
- Correlation coefficient analysis
- Taylor Diagrams

As defined by Karl E. Taylor (2001) provide a way of graphically summarizing how closely a pattern matches observations. The similarity between two patterns is quantified in terms of their correlation, their centered root-mean-square difference, and the amplitude of their variations

Neural Network Model and Results

General Flow



Data Set and Regressors For Training

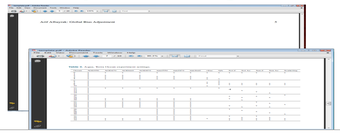
Data set for Terra-Aqua MODIS (ocean & land), is prepared by using Multisensor Aerosol Products Sampling System (MAPSS) database.

Training Data Set is prepared over a 10-year period from 2001 to 2011. The number of data Values used for training are summarized below:

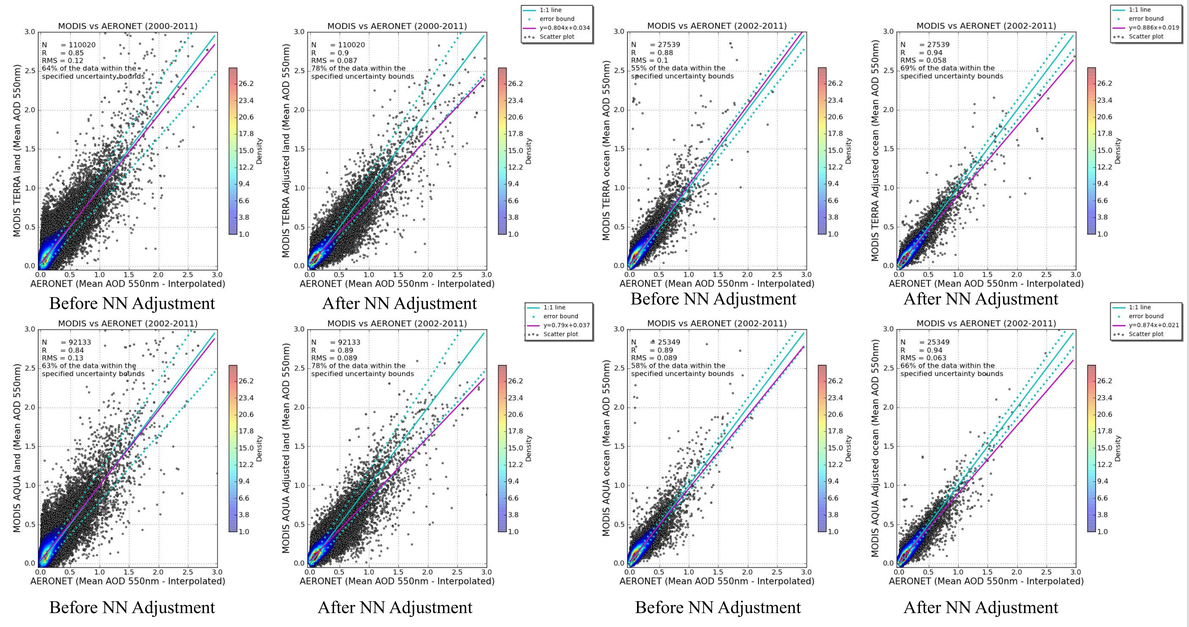
	Land	Ocean
Aqua	92133	25349
Terra	110020	27539

MODIS Common Regressors:
Target: mean AOD0550intrp (for training only)
Geometry: SolarZenith, SolarAzimuth, SensorZenith, SensorAzimuth, ScatteringAngle
Ocean: mean AOD0550, mean AOD0470, mean AOD0660, meanAOD0870, mean mref0550, meanmref0470, mean mref0660, mean cfrac, QAavg
Land: mean AOD0550, mean AOD0470, mean AOD0660, mean mref0470, mean mref0550, meansurfre0660, mean surfre0470, mean surfre2100, mean cfrac, QA

Preliminary Tests for Regressors



Bias Adjustment Results For AQUA and TERRA (Global)



Error Cones are the confidence intervals defined below:

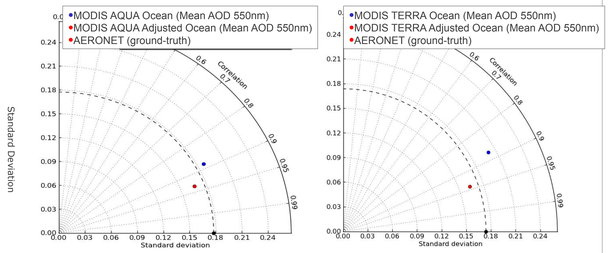
- Land = 0.15* Γ +/- 0.05
- Ocean = 0.05* Γ +/- 0.03
- The more the number of data in the cone, the better the results

Linear Fit and 1:1 line: The regression slope is affected by high AOD points while the majority of measurements are with low AOD. The NN approach treats both equally, so when the majority of the low AOD points are adjusted for their negative bias, the overall slope tilts down away from the 1:1.

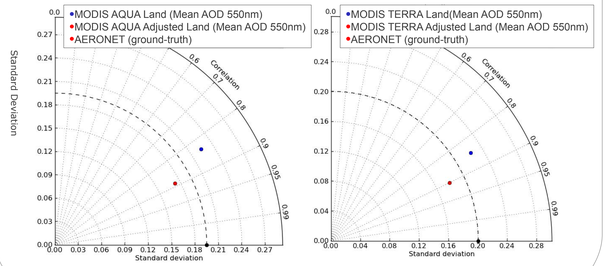
Summary

- Both correlation coefficient and standard deviation are improved after NN application.
- Independent analysis showed that there are 2 possible regimes in the data set causing the slope to move away from the 1:1 line. As a result, multiple statistics need to be used for realistic evaluation.
- The optimal regressor combination is currently studied by Dr. David Lary and his group.
- NN approach gives better approximations when using the Terra training data set than when using the Aqua data set for all cases.

Taylor Diagram (Ocean)



Taylor Diagram (Land)



Current and Future Work

Currently we are extending the work to Deep-Blue and MISR aerosol products. In the future we plan to:

- Consider multi regimes in the data sets: Run the same NN system on the global clusters that are obtained from multi-dimensional data sets.
- Combine regressors that are obtained from MISR and MODIS.
- Adjust biases for multiple data sets from different sources to a single reference and then merge them to extend coverage.

Acknowledgement

Support for the development of this on-line platform for statistical intercomparison of aerosols has been provided by NASA HQ (PM: Stephen Berrick) through the ROSES 2009 ACCESS Program (PI:Gregory Leptoukh). The AERONET data is a contribution by the International AERONET Federation (PI: Brent Holben).